Lab test use: what 1 billion claims tell us

written by CAP TODAY
August 19, 2024

Anne Paxton

August 2024—Scale back excessive laboratory testing and use the savings to test the undertested and to fund unreimbursed tests, such as molecular diagnostics, suggest the authors of a large-scale study that found significant overuse of four tests.

An analysis of nearly every outpatient HbA1c, PSA, vitamin D, and lipid panel test reimbursed by Medicare and most commercial payers in the U.S. in a single year might seem impossibly ambitious.

But researchers zeroing in on those four laboratory tests performed in 2019 carried out a retrospective study on precisely that scale, as they report in their recently published article, "Inappropriate Laboratory Testing: Significant Waste Quantified by a Large-Scale Year-Long Study of Medicare and Commercial Payer Reimbursement" (Smart D, et al. *Arch Pathol Lab Med*. Published online June 6, 2024. doi:10.5858/arpa.2023-0486-OA).

The study probed the level of low-value care—that is, care beyond what is required for managing health—in laboratory testing. Estimating the level of inappropriate laboratory testing has been far from an exact science, says coauthor Ila R. Singh, MD, PhD, professor of pathology and immunology at Baylor College of Medicine. Studying limited data sets, often from single institutions, makes it difficult to measure the scale of unnecessary testing because extrapolating estimates from small or narrow data sets to the entire population can lead to erroneous results.


Dr. Singh

In the latest study, Dr. Singh and coauthors addressed this problem on a national scale by examining approximately 1 billion claim records for outpatient tests from Medicare and most commercial payer databases. The claims covered all of 2019, a year that was chosen to avoid the confounding effects of the COVID-19 pandemic on test use. "We covered approximately 232 million individuals, and the U.S. adult population that year was 255 million, making our study quite comprehensive," Dr. Singh says.

"This was a groundbreaking study, and I'm not aware of anything else in laboratory medicine on this scale," she says. "With advanced computing power and very large data sets, we were finally able to analyze 1 billion claims." She credits coauthor Dave Smart, PhD, senior director of analytics at Diaceutics, based in Belfast, Northern Ireland, and his colleagues at the company with performing the data wizardry.

"This is truly big data. Any little change we made in the code, even a tweak," she says, required data queries to run the whole day and sometimes into the next. "We did that many, many, many times

because we were interested in looking at several features and in being as accurate as possible. The analysis took much longer than we thought it would."

The two main messages she hopes readers will take away: "One is that a lot of people are getting tested excessively. The other is that a lot of people are not getting tested at all."

The authors' analysis of paid outpatient test claim records for HbA1c, prostate-specific antigen, vitamin D 25-hydroxy, and lipid panel revealed that seven percent to 51 percent of the four tests they studied exceeded recommended frequencies—in some cases egregiously so, such as weekly HbA1c or PSA tests. The conservative cost estimate for these four excess tests was more than $350 million that year (using the 2019 clinical lab fee schedule). If the findings were extrapolated to all laboratory testing, the authors write, "Medicare alone may have incurred direct excess expenses from $1.95 to $3.28 billion in 2019, without factoring the hidden costs of excessive testing (eg, downstream care)."

The authors add, "Addressing unnecessary testing is crucial to lowering costs and redirecting resources for greater patient benefit."

Dr. Singh co-wrote the first Clinical and Laboratory Standards Institute document on test utilization management, a project spearheaded by Gary Procop, MD, MS. She, along with PLUGS (Patient-Centered Laboratory Utilization Guidance Services) and other collaborators, has played an important role in transforming the field initially called "laboratory utilization management" into "laboratory stewardship" and in establishing lab stewardship guidelines.

"I am delighted with the widespread adoption of the concept of laboratory test 'stewardship,' a term that has gained more acceptance than the previously used laboratory test 'utilization management,'" she says. "From our sister field of antimicrobial stewardship, we learned that people were more willing to be good 'stewards' rather than be 'managed' and told how much they had overutilized a test." Working on those papers with other leaders in stewardship led her to want to measure the true extent of the unnecessary testing "we were all trying to reduce," she says.

She also has worked to bring more clarity to laboratory test ordering. Dr. Singh is the founder of TRUU-Lab (Test Renaming for Understanding and Utilization in the Laboratory), a national initiative to standardize the names of clinical lab tests to make them easier to understand. TRUU-Lab is funded by the CDC, with the CMS and the FDA serving as liaisons, and with the participation of electronic health record vendors, the makers of instruments, and several professional organizations, including the CAP.

In her research now, Dr. Singh is leveraging AI and deep learning techniques to strengthen risk stratification, predictive analytics, health care resource use, and treatment protocols.

For the large-scale study on inappropriate lab testing, Dr. Singh and coauthors chose outpatient testing.

Many tests in this setting "have established guidelines," she says, "created by well-respected national organizations, that specify the appropriate frequency of testing."

They also considered CMS and commercial payer coverage guidelines and the published prevalence of relevant health and disease conditions.

The aims of this study, which is the first of a planned series, she says, were fairly simple despite its

scale. For example, to assess the frequency of HbA1c use, “in this first study we chose to look at whether an individual was tested for hemoglobin A1c. We chose two thresholds: one for people with diabetes and another for nondiabetics. Even though some recommendations say healthy nondiabetics should be tested only every two years, we wanted to err on the side of under- rather than overestimating excess testing, and decided that a yearly HbA1c was appropriate and not excessive.” The study found that 28 percent of individuals were tested twice or more, and 1.7 percent had five or more tests in 2019.

For all four tests in this study, “we are undercounting the excessive testing, not overcounting,” Dr. Singh says. “We are being tolerant in our [frequency] thresholds, and these excess costs we cite in the billions of dollars are underestimates, not overestimates.”

“As we point out in the paper,” Dr. Singh says, “physiologically, hemoglobin A1c is made by glucose binding to hemoglobin A irreversibly, and the amount of HbA1c reflects the amount of steady-state glucose you have in the body. The lifespan of red cells is 120 days, meaning your HbA1c today is about the same as it will be tomorrow.” Testing up to four times a year will provide more than adequate information. “And that frequency is recommended only for some diabetics. For diabetics with good glucose control, the recommendation is to measure HbA1c twice a year. So getting tested every month, every week for HbA1c is just a waste of money and time.”

It is possible that some clinicians lack a full understanding of recommended test frequencies. In other cases, the HbA1c test may be linked to other test orders and gets ordered inadvertently. Whatever the reason for overtesting, eliminating excess tests means more resources for the undertested, Dr. Singh suggests. According to Medicare’s data, only four percent of covered individuals are getting HbA1c tests for diabetes, even though, as this study shows, she says, there is enough money in the system to cover all those who are not getting tested.

Only about 10 percent of the population has vitamin D deficiency, though it varies by race, age, season, and other factors.

Since guidelines for the remaining 90 percent range from zero tests to one test per year, Dr. Singh and coauthors chose to consider one vitamin D test annually as an appropriate test frequency. They found 18.3 percent were tested twice or more per year. (An Endocrine Society guideline published in June advises against routine 25(OH)D testing in the absence of established indications: DeMay MB, et al. *J Clin Endocrinol Metab*. 2024;109[8]:1948–1954.)

Guidelines for PSA testing to screen for prostate cancer vary widely. Consistent with the intent to set tolerant thresholds, the study authors set one test per year as an acceptable test frequency for most men. Twenty percent were tested twice or more, 7.1 percent were tested more than three times, and 0.1 percent were tested monthly on average.

With lipid panels, the authors again took a more tolerant approach, considering one panel per year to be an acceptable frequency. Of the 25.78 million individuals tested in 2019, 23.5 percent were tested twice or more and 0.04 percent were tested at least 12 times in the year.

In their study population of 232 million, the authors found that about 14.4 million of about 60.5 million people (23.8 percent) had excessive lab testing.

Left out of the calculations are downstream costs, Dr. Singh says. One is the cost of an excess test

leading clinicians to order other tests. Some clinicians could be prompted by a marginally abnormal lab test result to order a series of other tests, she says. At other times, the patient may ask for further testing if a lab value is slightly out of range. “So when we test excessively, we make our patients vulnerable to further testing and even more costs.”

“We’re not counting the taxi ride to the doctor,” she adds. “We’re not counting the hours of lost work. We’re also not counting any ill effects arising from the phlebotomy or from unnecessary treatment.”

“While we are calling out the waste, we are also saying there are resources in the system for doing other things that are needed,” Dr. Singh continues. “In the world of utilization management and stewardship, there’s a common saying, ‘The pie is not getting any bigger,’ invoked when someone brings up the need for additional resources. What we’re doing with this study is creating more slices of the pie for things that might be more beneficial for our health than excessive tests.”

Their next study will examine tests that are performed less frequently than the guidelines recommend. “This is more challenging,” she says. “For example, in a study involving the *EGFR* mutation test, which assesses the likelihood of a lung cancer responding to certain anti-cancer drugs, we need to first verify if the patient has lung cancer and then measure how many people with lung cancer failed to receive the needed *EGFR* mutation test. We cite a study in our paper that showed that 64 percent of non-small cell lung cancer patients did not receive personalized therapy, at least in part because they did not get the appropriate test.”

Language can also play a part in appropriate test use, Dr. Singh says. What in the U.S. is called a test “order” is, in the U.K., Australia, and some other countries, called a test “request.”

“That makes a difference. As pathologists, we follow orders as they come. When we suspect an order is incorrect, our only recourse is to call the ordering physician and politely persuade them to change it. Although this approach usually works, it can be a long process. There simply aren’t enough pathologists to have these conversations each time a test is excessively ordered.”

“In countries where it’s called a test request, the pathologist feels empowered to change what appears to be incorrect. The way you name something can influence and transform the culture.”

Coauthor Jeff Schreier, BS, MBA, former senior director of partnerships for Diaceutics, became involved with the project about four years ago after meeting Dr. Singh at a conference and working with her on the TRUU-Lab initiative.

“In the mission to standardize test names, we were utilizing our data from Diaceutics to look at test volumes for specific analytes, namely vitamin D, when we made the discovery that there was a significant amount of over-ordering of that test. So we expanded the research to look at other analytes and ultimately settled upon four to study,” says Schreier, now founder and president of Molecupath Consultants, LLC.

Schreier

Part of his focus at Diaceutics and as an independent consultant has been on the case for redirecting existing funds in the payer system toward more high-value, critically needed testing. He views the *Archives* study's analysis of a billion claim records for outpatient testing as justifying a shift toward that way of thinking. "We hope the study will reinforce the need for providers and payers to redirect funds from low-value testing to more impactful high-value testing, especially precision medicine testing for cancer and rare diseases," he says.

Innovative precision medicine and companion diagnostics face unique challenges, Schreier says. "They follow regulatory pathways designed to get patients access to groundbreaking and potentially curative gene-based and cell-based therapies sooner." But the reimbursement system as it exists is not configured for these high-cost tests and therapies, especially when there are uncertain outcomes, he says. A key way to better allocate available health care system funds to the most effective uses, in his view, is to reduce the amount of unnecessary low-value testing. Each lab test is relatively inexpensive by itself but added up to billions of dollars in direct excess expenses to Medicare in 2019. "We're not saying, 'Do less testing.' We're saying do less testing that has little or no value. Use the funds to support testing that has high value in helping achieve beneficial patient outcomes," he says.

As the *Archives* study points out, artificial intelligence studies might be able to predict whether testing a patient for a particular analyte at a given time is useful, and Schreier sees great potential there. "AI and machine learning are already being put to use to generate real-world evidence to support reimbursement models in precision medicine testing that are tied to therapies for cancer and rare diseases." This makes a "realignment" in how health care funds are used important, he says.

The study found that excess testing among individuals covered by commercial payers was 30.5 percent of that seen in the CMS data.

"If you look at what level the commercial insurers stopped paying for tests, it's a little different from Medicare," Dr. Singh explains. "Their threshold is a little lower than Medicare's. But both allow for excessive levels of testing." She warns: "When addressing the issue of excessive testing, it's crucial that the costs don't get passed on to patients. The right thing to do is to simply avoid performing unnecessary tests."

A stewardship endeavor in the clinical laboratory is much harder to implement than in radiology, Dr. Singh says. "While avoiding a $3,000 CT scan is challenging, eliminating a $30 lab test performed 100 times is even more complex because of the need for 100 separate interventions. However, the sheer volumes of these lab tests is why we must pay attention to even seemingly inexpensive tests."

Some might fear the impact on laboratories if testing is reduced through stewardship programs. But the resources saved, she says, will enable reimbursement for much testing that's currently not covered. "Many labs developing cutting-edge testing technologies in academic hospitals, such as molecular pathology labs, are currently underfunded. Better resourcing of these labs will allow many future pathology jobs to emerge."

Anne Paxton is a writer and attorney in Seattle.